第 52 卷 第 6 期：2729-2737 高电压技术 Vol.52, No.6: 2729-2737
2026 年 6 月 30 日 High Voltage Engineering June 30, 2026
DOI: 10.13336/j.1003-6520.hve.20251068

基于 WGAN 数据增强和 ISSA-SVM 的变压器故障诊断

曾贤强，梁　嘉，刘仲民，李恒杰
（兰州理工大学自动化与电气工程学院，兰州 730050）

摘　要：针对变压器故障样本不均衡以及故障模型易陷入局部最优，从而影响故障识别性能的问题，提出一种融合基于 Wasserstein 距离生成对抗网络(Wasserstein generative adversarial network，WGAN)与改进麻雀搜索算法(improved sparrow search algorithm，ISSA)优化支持向量机(support vector machine，SVM)的变压器故障诊断方法。首先，利用 WGAN 对样本数据进行增强，缓解样本不均衡。其次，引入 Logistic 混沌映射、随机游走策略与自适应萤火虫算法对 SSA 进行改进，以提升其参数寻优能力，并用于 SVM 参数优化，构建 ISSA-SVM 故障诊断模型。最后，在基于 WGAN 增强后的数据集上进行故障诊断实验。实验结果表明，该方法在样本不均衡条件下仍具备较高的诊断性能，诊断准确率达到 96.60%，优于 SSA-SVM 模型及其他数据平衡方法。
关键词：变压器；故障诊断；样本不均衡；WGAN；改进麻雀搜索算法；支持向量机

Transformer Fault Diagnosis Based on WGAN Data Augmentation and ISSA-SVM

ZENG Xianqiang, LIANG Jia, LIU Zhongmin, LI Hengjie
(School of Automation and Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: The imbalanced transformer fault samples and the tendency of fault diagnosis models to fall into local optima may degrade fault identification performance. To address the issues above, this paper proposes a transformer fault diagnosis method that integrates a Wasserstein generative adversarial network(WGAN) with a support vector machine(SVM) optimized by an improved sparrow search algorithm(ISSA). First, a WGAN-based data augmentation strategy is employed to expand the sample set and mitigate class imbalance. Second, Logistic chaotic mapping, a random walk strategy, and an adaptive firefly algorithm are introduced to improve the standard sparrow search algorithm(SSA), thereby enhancing its parameter optimization capability. The improved algorithm is then used to optimize the key parameters of SVM, and an ISSA-SVM fault diagnosis model is established. Finally, fault diagnosis experiments are conducted on the WGAN-augmented dataset. The experimental results show that the proposed method maintains high diagnostic performance under imbalanced sample conditions, achieving a diagnostic accuracy of 96.60%, outperforming the SSA-SVM model and other data balancing methods.
Key words: transformer; fault diagnosis; class imbalance; WGAN; improved sparrow search algorithm; support vector machine

0　引言

电力变压器是配电系统中的核心设备，其运行状态直接关系到电力系统的安全与稳定。油浸式变压器在运行过程中一旦发生内部故障，往往会伴随特征气体溶解于绝缘油中。研究表明，油中溶解气体的含量与故障类型之间存在显著关联。基于此，油中溶解气体分析(dissolved gas analysis，DGA)已成为诊断油浸式电力变压器内部故障的重要技术手段[1-2]。传统 DGA 方法如 Rogers 比值法、IEC 三比值法等依赖于专家经验和固定规则，在样本模糊、故障类型交叉等复杂场景下，诊断精度和泛化能力均存在局限[3-4]。随着人工智能的发展[5]，基于数据驱动的机器学习方法被广泛应用于 DGA 数据分析

基金资助项目：国家自然科学基金(52167014)；甘肃省重点研发项目(23YFGA0084)；甘肃省联合科研基金项目(24JRRA827)；甘肃省中央引导地方科技发展资金(25ZYJN001)；甘肃省重点研发计划-工业类(25YFGA020)。
Project supported by National Natural Science Foundation of China (52167014), Key R&D Project of Gansu Province(23YFGA0084), Joint Research Fund Project of Gansu Province(24JRRA827), Cetral-Guided Local S&T Development Fund of Gansu Province(25ZYJN001), Key R&D Program of Gansu Province-Industrial Category(25YFGA020).

中[6-7]，例如人工神经网络[8]、极限学习机[9]以及支持向量机[10]等。其中，SVM 凭借良好的泛化能力与在小样本条件下的优越表现，已成为故障识别领域的重要工具[11-12]。但是SVM的分类性能对参数敏感，如何实现高效、稳定的参数优化仍是研究难点[13]。

近年来，群体智能优化算法在模型参数寻优中表现出较强的优势。麻雀搜索算法(sparrow search algorithm，SSA)作为新兴算法[14]，因其收敛速度快、结构简单等特点受到广泛关注。其中，文献[15]通过引入改进 Tent 混沌映射对 SSA 种群初始化，提高了种群的初始化。文献[16]提出了两阶段SSA(two-stage SSA，TS-SSA)，采用自适应种群划分及多种搜索机制等分别处理收敛性和多样性。文献[17]将差分进化策略融入 SSA 中(modified SSA，MSSA)，提升了算法的全局搜索能力和稳定性。以上文献对 SSA 算法进行改进，在一定程度上提升了其性能，但在搜索精度和跳出局部最优方面仍存在不足。

此外，DGA 样本的不均衡分布也是制约诊断模型性能的关键因素。由于故障样本获取困难，导致数据集类别分布不均衡，直接训练机器学习模型易使模型偏向多数类样本，从而影响分类精度。当前常用的样本平衡技术有欠采样和过采样。欠采样通过减少多数类样本数量实现类别均衡，虽能在一定程度上缓解数据不平衡问题，但可能导致原始数据中关键信息的丢失。传统过采样方法(如 SMOTE)通常基于线性插值生成新样本，忽略了类别边界与样本分布结构，合成的样本质量有限，难以提升模型的故障识别能力[18-19]。

为此，本文提出一种融合 Wasserstein 生成对抗网络与改进麻雀搜索算法优化支持向量机的变压器故障诊断方法。该方法首先利用 WGAN 对原始数据集进行数据增强，提升数据多样性与完整性，随后引入融合 Logistic 混沌映射、随机游走策略与自适应萤火虫算法的改进 SSA 算法对 SVM 模型参数进行优化，从而实现对变压器故障类型的高精度识别。该方法在保证样本完整性与多样性的同时，有效提升了诊断模型的全局搜索能力与分类精度，为油浸式变压器故障诊断提供了一种高效可行的新方案。

1 WGAN 数据增强

1.1 WGAN 模型

传统过采样方法在处理类别不平衡问题时，往往难以生成具有充分多样性的样本，限制了模型对少数类的识别能力。生成对抗网络作为一种基于深度学习的生成模型[20]，能够有效学习数据分布并生成高质量的合成样本，在应对样本不平衡问题中展现出显著优势。然而，经典 GAN 采用 Jensen-Shannon(JS)散度作为衡量生成分布与真实分布差异的指标，当两者分布差异较大时，JS 散度的梯度趋近于零，易导致梯度消失，从而影响训练的稳定性，甚至引发模式崩塌现象[21]。WGAN 通过引入 Wasserstein 距离替代 JS 散度，有效缓解了上述问题，显著提升了模型训练的稳定性与收敛性，并增强了生成样本的质量和多样性。其 Wasserstein 距离定义为：

$$W\left(P_{\mathrm{r}},P_{\mathrm{g}}\right)=\inf_{y\sim\prod\left(P_{\mathrm{r}},P_{\mathrm{g}}\right)}\left\{E_{(x,y)\sim y}\left(\left\|x-y\right\|\right)\right\} \tag{1}$$

式中：$W(P_{\mathrm{r}},P_{\mathrm{g}})$为真实样本分布 P_{r} 和生成样本分布 P_{g} 之间的 Wasserstein 距离；x 是真实样本；y 是生成样本；$\Pi(P_{\mathrm{r}},P_g)$是 P_{r} 与 P_{g} 组合起来的所有可能的联合分布的集合，y~Π(Pr,Pg)表示该集合中的一种具体联合分布，即真实样本与生成样本之间的一种可行匹配关系；$E_{(x,y)\sim y}(\|x-y\|)$是样本距离$\|x-y\|$的期望；inf{•}表示集合的下确界。

为将 Wasserstein 距离应用于生成模型的训练中，WGAN 引入 Kantorovich-Rubinstein 对偶性，将原始的最优传输问题转化为一个可优化的形式。具体地，判别器(critic)不再输出样本为真实的概率值，而是学习一个满足 1-Lipschitz 条件的实值函数，以度量生成分布与真实分布之间的差异。基于此，WGAN 的目标函数可表示为：

$$\min_{G}\max_{D}E_{x\sim P_{r(x)}}\left(D\left(x\right)\right)+E_{z\sim P_{z(z)}}\left(D\left(G\left(z\right)\right)\right) \tag{2}$$

式中：$P_r(x)$表示真实样本数据分布；E 表示数学期望；G 表示生成器；D 表示判别器；x 是真实样本；z 是从高斯分布 $P_z(z)$中随机采样的噪声向量；$x\sim P_r(x)$表示 x 服从分布 $P_r(x)$；$z\sim P_z(z)$表示随机噪声 z 服从高斯分布 $P_z(z)$。

1.2 设备运行状态编码及模型参数

本文选取 H_2、CH_4、C_2H_6、C_2H_4 和 C_2H_2 气体含量作为变压器的故障特征[22]，可将电力变压器运行状态划分为正常、中低温过热、高温过热、低能放电、局部放电和高能放电 6 种状态，利用 One-Hot 编码对各种运行状态进行编码[23]，其结果见表 1。

采用 WGAN 生成数据量较少的类别数据。WGAN 中优化器均采用 RMSProp 优化器，学习率

设置 5×10^{-5}，生成器由多层一维卷积神经网络(Conv1d)、批量归一化层以及上采样层组成，隐含层的激活函数为 LeakyReLU，输出层的激活函数为 Tanh，Conv1d 中的步长(stride)和填充(padding)均设置为 1，判别器均使用 3 层全连接层，隐含层的激活函数为 LeakyReLU，输出层去掉 Sigmoid 激活函数，其模型网络结构如表 2 所示。

表 1　One-Hot 编码结果

Table 1　One-Hot encoding results

运行状态	数字编码	One-Hot 编码
中低温过热	1	(1,0,0,0,0,0)
高温过热	2	(0,1,0,0,0,0)
局部放电	3	(0,0,1,0,0,0)
低能放电	4	(0,0,0,1,0,0)
高能放电	5	(0,0,0,0,1,0)
正常	6	(0,0,0,0,0,1)

表 2　WGAN 模型网络结构

Table 2　WGAN model architecture

网络	结构	激活函数
生成器	Linear(***Z***→128×初始长度)+Reshape	None
	BatchNorm1d(128)	None
	Upsample(scale_factor=2)	None
	Conv1d(128,128,3, stride=1, padding=1)+BatchNorm1d	LeakyReLU
	Upsample(scale_factor=2)	None
	Conv1d(128,64,3, stride=1, padding=1)+BatchNorm1d	LeakyReLU
	Conv1d(128, img,3, stride=1, padding=1)+BatchNorm1d	Tanh
判别器	Linear(输入维度→512)	LeakyReLU
	Linear(512→256)	LeakyReLU
	Linear(256→128)	LeakyReLU
	Linear(128→1)	None

1.3　增强数据集

为解决样本类别分布不均衡问题，本文采用 WGAN 数据增强方法对样本类别进行平衡化处理，WGAN 结构如图 1，包括生成器(Generator)与判别器(Discriminator)两个核心模块。

图 1 中，在生成器部分，首先将服从高斯分布的随机噪声向量 $\boldsymbol{Z}\in\mathbf{R}^d$ 输入至生成器网络，经全连接层映射为维度 $128L/4$ 的高维特征向量，其中 L 为目标样本的长度维度。随后，通过形状重塑(Reshape)操作将该特征向量重构为 $128\times L/4$ 的一维多通道特征图，以适配后续一维卷积操作。重构后的特征依次通过批归一化(BatchNormalization)、非线性激活函数(LeakyReLU)、上采样(Upsampling)与一维卷积神经网络操作，逐步完成特征增强与分辨率恢复，其中，上采样模块采用尺度因子(scale_factor)为 2，即每次上采样将特征序列的长度维度扩大为原来的 2 倍，而通道数保持不变。经过两次上采样后，特征长度由 $L/4$ 逐步恢复至 L，最终生成与真实样本尺寸一致的假样本 X_{fake}，生成器末端采用 Tanh 激活函数，将输出范围压缩至[−1,1]，与归一化后的真实样本分布一致；在判别器部分，输入图像(无论是真实样本或生成样本)首先通过展平操作转换为一维向量，并依次输入 4 层全连接网络(Linear1—Linear4)以提取判别特征。各层间采用 LeakyReLU 激活函数以增强非线性表达能力，最终输出单个实值，以衡量输入图像的真实性。训练过程中，判别器对生成图像进行识别反馈，并通过 Wasserstein 损失函数对生成器进行反向优化，以实现更稳定的训练过程。具体过程如下。

以中低温过热样本为例，利用式(2)计算该类别的生成器和判别器的最优参数 θ_{G} 和 θ_{D}，使得损失函数 L_{G}、L_{D} 达到最小。为了使生成的样本分布尽可能接近真实样本分布，采用判别器训练 5 次，生成器训练 1 次的方式。过多轮对抗训练，生成器和判别器逐步达到纳什平衡，从而生成高质量假样本 X_{fake}，其表达式如下：

$$X_{\text{fake}}=G\left(z;\theta_{\text{G}}\right) \tag{3}$$

式中：X_{fake} 表示生成的假样本。

将以上生成的假样本与真实样本混合，来增强该类别的数据集。同样，对之后新增的故障样本重复此步骤，以此增强所有类别的数据集。

2　基于 WGAN 数据增强的 ISSA-SVM 变压器故障诊断

2.1　麻雀搜索算法

麻雀搜索算法(sparrow search algorithm，SSA)作为一种新兴的群体智能优化方法，具有收敛速度快、结构简单等优点。然而，原始 SSA 在种群初始化阶段采用均匀随机初始化方式，可能导致初始种群的分布不均，进而影响全局搜索能力，并在后期容易陷入局部最优解，导致搜索精度不足。针对该问题，本文对 SSA 进行了以下 3 方面的改进：(1)引入 Logistic 混沌映射进行初始化，以提升初始种群的多样性和覆盖能力；(2)通过对当前最优个体引入随机游走扰动，增强算法的全局搜索能力；(3)结合自适应萤火虫算法对种群位置进行精细化扰动，从而提高局部搜索的精度和灵活性。

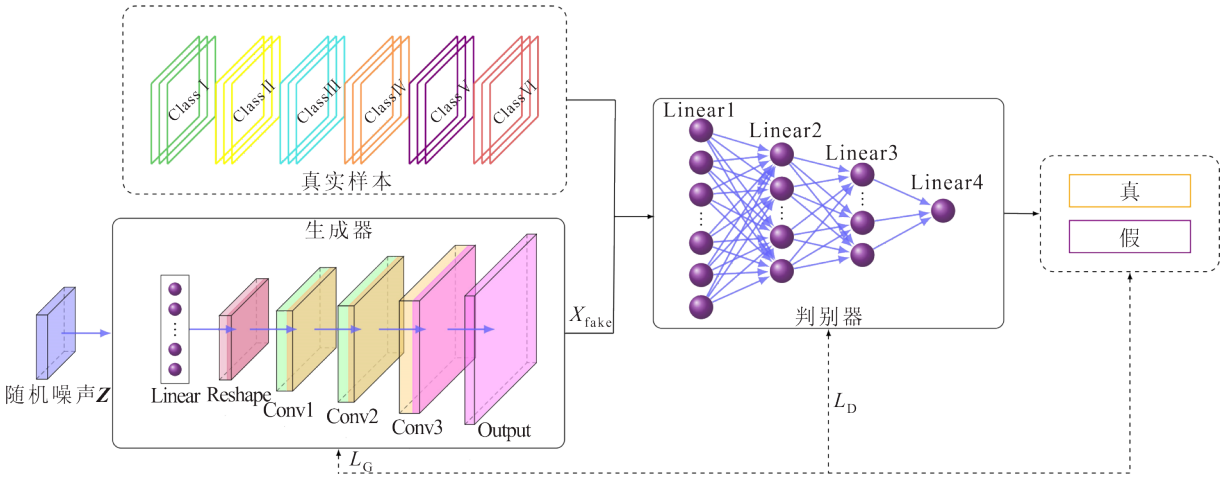


图 1 WGAN 模型网络结构图

Fig.1 Network architecture diagram of the WGAN model

2.1.1 Logistic 混沌映射

Logistic 映射具有初值敏感和遍历性强的混沌特性，能有效提升初始种群在搜索空间内的分布均匀性[24]，其表达式如下：

$$x_{n+1}=ax_n\left(1-x_n\right) \tag{4}$$

式中：x_n 表示第 n 代的混沌值，$x_n\in(0,1)$；a 是系统参数，可控制系统的动态行为，决定了系统是呈现周期性行为、混沌行为还是稳定状态，当 a=4 时系统呈现完全混沌状态，生成序列具有遍历性和类随机性；x_{n+1} 是第 n+1 次迭代的状态值。

在 SSA 中，个体的初始位置一般通过如下公式均匀随机生成：

$$X_{i,j}=X_{\min,j}+rand\cdot\left(X_{\max,j}-X_{\min,j}\right) \tag{5}$$

式中：$X_{i,j}$表示第 i 个体第 j 个维度的初始位置；$X_{\min,j}$、$X_{\max,j}$分别表示第 j 个维度的下限与上限；$rand$ 为(0,1)之间的任意随机数。

为提升 SSA 种群初始化的多样性与覆盖性，引入 Logistic 混沌映射进行个体位置初始化。该方法通过 Logistic 映射生成分布于[0,1]区间的混沌序列，并将其映射至变量的搜索空间，从而构建具有更强遍历性的初始种群。

改进后的位置初始化计算式如下：

$$X_{i,j}=X_{\min,j}+x_n\cdot\left(X_{\max,j}-X_{\min,j}\right) \tag{6}$$

2.1.2 随机游走策略

随着 SSA 的迭代进行，种群差异性逐渐减弱，个体位置趋于一致，尤其在后期更容易陷入局部最优区域。为增强算法的全局搜索能力，本文引入随机游走策略。该策略通过基于当前解的位置进行扰动更新，在每次迭代中对发现者进行扰动，帮助跳出局部最优区域[25]。随机游走的表示式如下：

$$X(t)=\left[0,cussum\left(2r\left(t_1\right)-1\right),\cdots,cussum\left(2r\left(t_n\right)-1\right)\right] \tag{7}$$

式中：$X(t)$为随机游走的步集数；$cussum$ 为计算累加和；t 为随机游走的步数；t_n 表示随机游走的最大步数；$r(t)$为二值随机函数，其定义为：

$$r(t)=\begin{cases}1, & rand>0.5\\ 0, & rand\le 0.5\end{cases} \tag{8}$$

由于可行域存在边界，不能直接使用式(7)更新麻雀的位置。为确保行走在可行域范围内随机游走，需根据式(9)对其进行归一化：

$$\tilde{X}_i^t=\frac{\left(X_i^t-a_i\right)\cdot\left(d_i^t-c_i^t\right)}{\left(b_i-a_i\right)}+c_i^t \tag{9}$$

式中：X_i^t 为第 t 次迭代中的当前发现者位置；a_i 为第 i 维变量随机游走的最小值；b_i 为第 i 维变量随机游走的最大值；c_i^t 为第 i 维变量在第 t 次迭代的最小值；d_i^t 为第 i 维变量在第 t 次迭代最大值。

通过引入随机扰动后，SSA 的发现者位置更新公式为：

$$X_i^{t+1}=X_i^t+\beta\cdot\left(X_{\text{best}}^t-X_i^t\right)+r'(t)\cdot R(t) \tag{10}$$

式中：X_i^{t+1} 为更新后的发现者位置；X_{best}^t 为当前全局最优解；β 是缩放因子，用于控制发现者向全局最优解逼近的速度；$r'(t)$ 为扰动幅度，随时间变化逐渐减小；$R(t)$是通过随机游走生成的扰动项。

2.1.3 自适应萤火虫算法

原始 SSA 通过“发现者—追随者—警戒者”的角色划分更新个体位置，但在局部搜索精度和跳出局部极值能力方面仍有不足。为了增强 SSA 的全局

搜索能力并提升收敛精度，本文引入了一种自适应扰动机制的萤火虫算法(adaptive firefly algorithm, AFA)对 SSA 进行协同优化[26]。该方法在每一轮 SSA 位置更新后，引入基于距离吸引模型和扰动因子的精细化搜索策略，从而进一步提高跳出局部最优解的能力。自适应萤火虫位置更新公式如下：

$$\tilde{X}_i^{t+1}=\tilde{X}_i^t+\lambda\cdot\left(\tilde{X}_j^t-X_i^t\right)+\alpha\cdot(\varepsilon-0.5) \quad (11)$$

式中：$\tilde{X}_i^t$ 表示第 t 代中第 i 个个体的位置；$\tilde{X}_j^t$ 是吸引源个体，此处选取当前最优个体 X_{best}^t；α 为自适应扰动步长，采用指数衰减动态策略调整，即 $\alpha=\max(\alpha_0\rho^t,\alpha_{\min})$，$\alpha_0$ 为初始扰动步长，$\alpha_{\min}$ 为最小扰动步长，$\rho\in(0,1)$为衰减因子；$\varepsilon\sim U(0,1)^d$ 为服从均匀分布的随机向量，d 为搜索空间维度；λ 为吸引度因子，其随个体距离变化而衰减，数学表达式为：

$$\lambda=\lambda_0\cdot\exp\left(-\gamma\cdot r_{i,\text{best}}^2\right) \quad (12)$$

式中：γ 表示光强吸收因子；$r_{i,\text{best}}$ 为个体与最优个体间的归一化欧几里得距离；λ_0 为最大吸引度。

为简化混合算法框架，将 SSA 的标准位置更新表示为：

$$X_i^{\text{SSA}}=F_{\text{SSA}}\left(X_i^t\right) \quad (13)$$

式中：$F_{\text{SSA}}(\bullet)$表示根据个体角色(发现者/加入者/警觉者)所采取的 SSA 更新算子。

在 SSA 完成一次位置更新后，引入 AFA 扰动生成新候选位置：

$$X_i'=X_i^{\text{SSA}}+\beta\cdot\left(X_i^{\text{SSA}}-X_{\text{best}}^t\right)+\alpha\cdot(\varepsilon-0.5) \quad (14)$$

基于上述模型，最终的更新策略可以表示为：

$$X_i^{t+1}=\begin{cases}X_i', & f\left(X_i'\right)<f\left(X_i^{\text{SSA}}\right)\\ X_i^{\text{SSA}}, & \text{其他}\end{cases} \quad (15)$$

即在 SSA 更新后，个体进一步通过 AFA 扰动获取新位置，若其适应度优于原位置，则进行更新；否则保持不变。通过上述融合策略，保留了 SSA 的全局搜索能力，同时利用 AFA 进行局部扰动增强，从而实现全局与局部协同优化，提升整体搜索性能。

2.2　WGAN-ISSA-SVM 模型构建

本文提出的基于 WGAN 数据增强的 ISSA-SVM 油浸式变压器故障诊断模型包括 WGAN 模型进行数据增强、基于 ISSA 算法的 SVM 模型参数优化、基于 ISSA-SVM 模型的变压器故障诊断 3 个阶段。模型的故障诊断流程如图 2 所示。

其具体诊断流程步骤如下：

1）将 DGA 数据进行归一化并输入到 WGAN 模型中，对生成器和判别器进行交替对抗训练，优化样本分布，使生成样本更贴近真实数据。

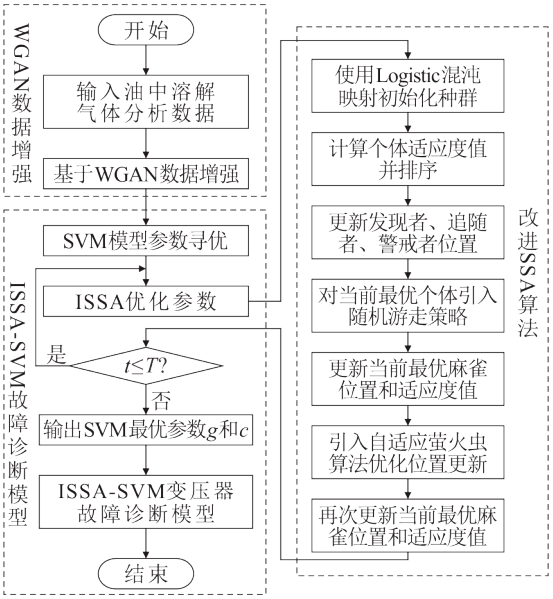


图 2　基于 WGAN-ISSA-SVM 故障诊断流程图

Fig.2　Flowchart of the fault diagnosis process based on WGAN - ISSA-SVM

2）利用 Logistic 混沌映射初始化搜索空间内种群个体位置，提升初始种群的多样性与全局覆盖能力，每个个体代表 SVM 中的核函数参数 g 和惩罚因子 c。

3）计算每个个体的适应度并排序，确定当前全局最优个体。根据 SSA 的标准位置更新规则，分别更新发现者、追随者、警戒者位置。

4）对当前全局最优解引入式(7)进行随机游走扰动，以提高跳出局部最优的能力，若扰动后新个体的适应度优于原全局最优个体，则更新全局最优位置与适应度值。

5）进一步对全部个体引入式(11)的自适应萤火虫算法扰动机制，以增强局部搜索能力与收敛精度，并对扰动后的个体再次更新评估适应度，若更优则替代原解进入下一代。

6）判断当前迭代次数 t 是否达到最大迭代次数 T，若满足条件，则输出最优参数 g 和 c，否则返回 2）继续迭代，最后构建 ISSA-SVM 变压器故障诊断模型。

7）将 WGAN 生成的样本数据与原始数据混合在一起，并按 8:2 划分其测试集和训练集，将训练集输入到 ISSA-SVM 变压器故障诊断模型进行诊断，测试集进行验证，最终输出故障诊断结果。

3 变压器故障诊断实例

3.1 数据预处理

本文从公开发表的文献中收集到 255 组油中溶解气体含量数据[27]，涵盖故障和正常状态。首先对原始数据进行归一化处理，使其缩放至[0,1]区间，以消除不同特征之间的量纲影响。随后将归一化后的数据输入 WGAN 模型进行训练并生成样本，将生成样本与原始样本混合，最终构建出包含 1 623 组样本的增强数据集。将增强后的数据集按 8:2 比例划分为训练集与测试集，分别包含 1 329 组与 294 组样本。利用训练集对改进的 ISSA-SVM 模型进行训练，并采用准确率($A_{ccuracy}$)、精确率($P_{recision}$)、召回率(R_{ecall})以及 F_1 值对模型性能进行评估。

图 3 及表 3 展示了基于 WGAN 数据增强的 ISSA-SVM 模型的混淆矩阵与各故障类型的性能指标。可以看出，本文所使用的方法其故障诊断准确率高达 96.60%，且各故障类型的精确率、召回率及 F_1 分数均高于 92%，体现了本文所使用的模型在多分类任务中具有显著的优势。

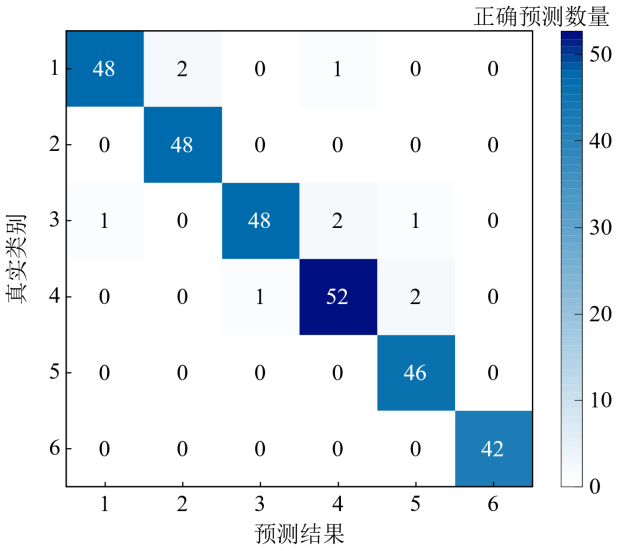


图 3 故障诊断结果混淆矩阵

Fig.3 Confusion matrix of the fault diagnosis results

3.2 ISSA 寻优性能测试

为验证改进麻雀搜索算法的寻优性能，使用 Ackley 标准测试函数对 PSO、GA、SSA 以及 ISSA 进行函数测试。其中 Ackley 函数表达式为：

$$f(x)=-20\exp\left(-0.2\sqrt{\frac{1}{30}\sum_{i=1}^{30}x_i^2}\right)-\exp\left(\frac{1}{30}\sum_{i=1}^{30}\cos(2\pi x_i)\right)+20+\mathrm{e} \tag{16}$$

表 3 WGAN-ISSA-SVM 模型评价指标

Table 3 Performance evaluation metrics for the WGAN-ISSA-SVM

故障编码	评价指标			
	精确率/%	召回率/%	F_1/%	准确率/%
1	97.96	94.12	96.00	96.60
2	96.00	100.00	97.96	
3	97.96	92.31	95.05	
4	94.55	94.55	94.55	
5	93.88	100.00	96.84	
6	100.00	100.00	100.00	

式中：x_i 表示第 i 维决策变量；e 为自然常数。

根据图 4 可得，在 Ackley 基准函数测试过程中，GA 与 PSO 算法收敛速度最慢，SSA 算法在第 152 次迭代时开始收敛，而 ISSA 算法在迭代 72 次时就开始收敛。

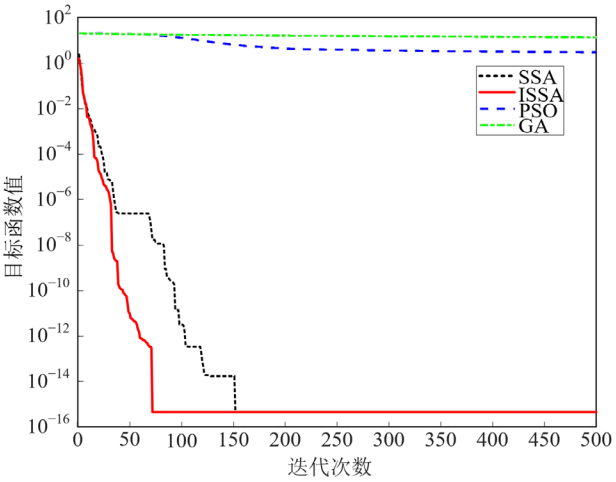


图 4 不同优化算法对比图

Fig.4 Comparison of different optimozation algorithms

3.3 不同数据平衡方法对比

本文使用 WGAN 进行数据增强，并与 SMOTE、GAN 以及原始数据集方法进行比较。图 5—图 8 分别为 4 种数据处理方式下基于 ISSA-SVM 的预测结果与实际结果对比图。分析不同平衡处理方法结果诊断精度图可知，本文提出的引入 WGAN 模型进行数据增强，有效地解决了故障样本不平衡问题，其故障诊断率效果最优，达到 96.60%。使用原始数据集的 ISSA-SVM 模型诊断精度为 92.73%，提升了 3.87%。与 SMOTE 算法以及 GAN 平衡化方法相比，分别提升了 4.45%、2.72%。结果证实，本文所提出的 WGAN 进行数据增强后的模型的故障诊断具有显著的优势。

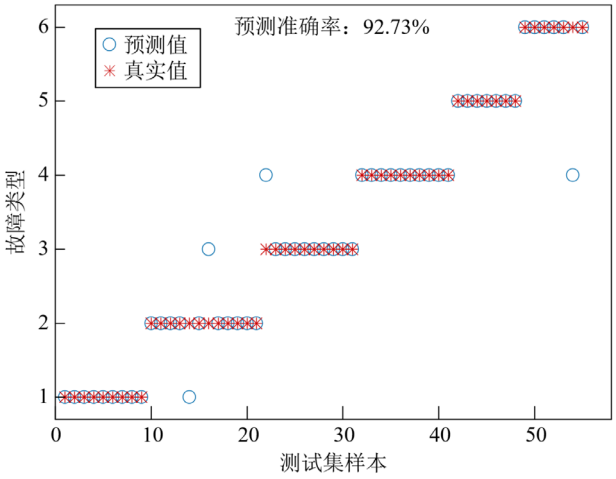


图 5　基于原始数据集的 ISSA-SVM 诊断精度图

Fig.5　Diagnostic accuracy of ISSA-SVM based on the original dataset

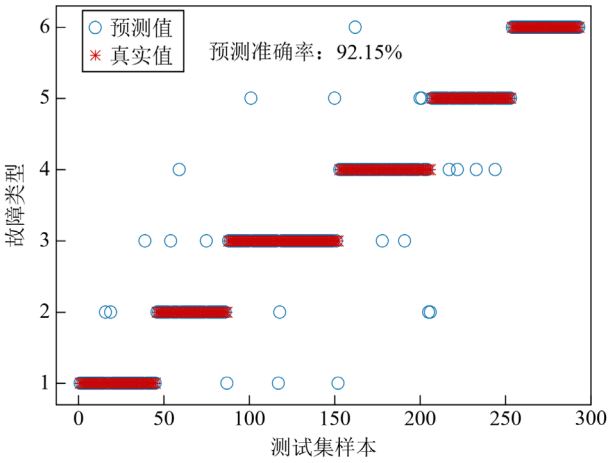


图 6　基于 SMOTE 数据扩充的 ISSA-SVM 诊断精度图

Fig.6　Diagnostic accuracy of ISSA-SVM based on SMOTE-augmented data

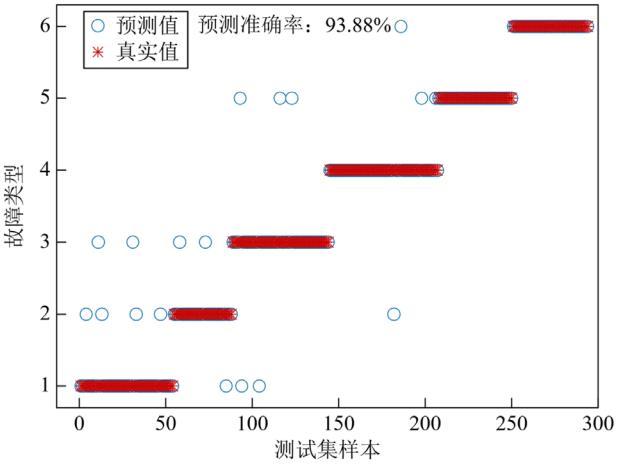


图 7　基于 GAN 数据增强的 ISSA-SVM 诊断精度图

Fig.7　Diagnostic accuracy of ISSA-SVM based on GAN-augmented data

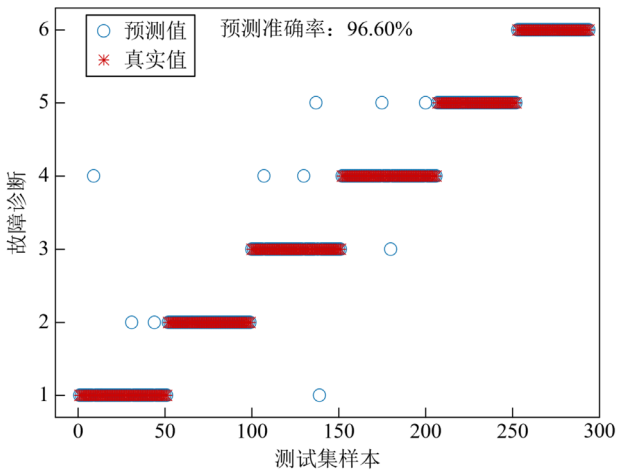


图 8　基于 WGAN 数据增强的 ISSA-SVM 诊断精度图

Fig.8　Diagnostic accuracy of ISSA-SVM based on WGAN-augmented data

3.4　模型的综合比较

为验证所提的 WGAN-ISSA-SVM 故障诊断方法的有效性，本文将其与 SSA-SVM、ISSA-SVM、SMOTE-SSA-SVM 以及 GAN-ISSA-SVM 模型进行对比。综合比较结果如表 4 所示。由表 4 可知，基础 SVM 模型的诊断准确率仅为 70.37%，在引入 SSA 算法后，模型准确率提升至 88.89%，提高了 18.52%，说明 SSA 在 SVM 参数优化方面具有显著效果。进一步引入改进的 ISSA 算法后，模型准确率提升至 92.73%，相较于 SSA 提升了 3.84%，表明改进后的算法不仅提升了搜索能力，还有效避免陷入局部最优的问题，优化后的模型具有更强的全局搜索能力和更快的收敛速度。在不同改进模型的

表 4　模型的综合比较结果

Table 4　Comprehensive comparison results of the models

分类模型	准确率/%	精确率/%	召回率/%	F_1/%
SVM	70.37	72.63	68.89	68.21
PSO-SVM	85.19	85.05	85.94	84.93
GA-SVM	87.04	88.27	87.04	86.99
SSA-SVM	88.89	90.91	89.72	89.68
TS-SSA-SVM	88.89	89.08	89.03	88.64
MSSA-SVM	90.74	92.14	88.96	88.95
ISSA-SVM	92.73	93.89	93.17	93.14
SMOTE-ISSA-SVM	92.15	92.24	92.44	92.33
GAN-ISSA-SVM	93.88	93.39	94.13	93.67
WGAN-ISSA-SVM	96.60	96.72	96.83	96.73

比较中，WGAN-ISSA-SVM 模型的综合性能最优，其准确率达到 96.60%，精确率、召回率和 F_1 值分别为 96.72%、96.83%和 96.73%，均高于其他对比模型，进一步验证了本文所提出方法在故障诊断任务中的有效性和优越性。

4 结论

针对变压器故障样本不平衡及故障识别模型易陷入局部最优的问题，提出了一种基于 WGAN 增强与改进麻雀搜索算法(ISSA)优化支持向量机(SVM)的故障诊断方法。相关实验验证了所提方法的有效性，结论如下：

1）在原始数据集下，所构建的 ISSA-SVM 模型在准确率、精确率、召回率及 F_1 值等关键性能指标上，均优于 PSO-SVM、GA-SVM、MSSA-SVM、TS-SSA-SVM 和原始 SSA-SVM 等对比模型，表明所提 ISSA 方法在支持向量机参数寻优方面具备更优的收敛能力和泛化性能。

2）在 ISSA-SVM 模型基础上引入 WGAN 进行数据增强后，模型性能进一步提升，准确率、精确率、召回率和 F_1 值分别提高了 3.87%、2.83%、3.66%和 3.59%。相较于传统过采样 SMOTE 和经典 GAN 常用增强策略，本文方法在各项指标上均取得最优结果，验证了其在处理故障样本不平衡问题中的适应性与稳定性。

3）所提出的基于 WGAN 的样本增强方法有效缓解了故障样本不均衡问题，提高了模型在少数类识别中的准确性；同时，通过引入 Logistic 混沌映射、随机游走策略与自适应萤火虫算法，对 SSA 算法进行了有效改进，显著提升了模型的全局搜索能力与收敛效率。实验结果表明，WGAN 与 ISSA 的融合策略在提升变压器故障诊断精度与计算效率方面表现出良好的工程适应性和应用前景。

参考文献 References

[1] AL-SAKINI S R, BILAL G A, SADIQ A T, et al. Dissolved gas analysis for fault prediction in power transformers using machine learning techniques[J]. Applied Sciences, 2024, 15(1): 118.

[2] 廖才波，杨金鑫，邱志斌，等. 基于缺失数据填补的油浸式变压器故障诊断[J]. 高电压技术，2024，50(9)：4091-4100.
LIAO Caibo, YANG Jinxin, QIU Zhibin, et al. Fault diagnosis for oil-immersed transformer based on missing data imputation[J]. High Voltage Engineering, 2024, 50(9): 4091-4100.

[3] 周萱，吴伟丽. 基于改进 SMOTE 不均衡样本处理和 IHPO-DBN 的变压器故障诊断方法研究[J]. 电力系统保护与控制，2024，52(11)：21-30.
ZHOU Xuan, WU Weili. Transformer fault diagnosis method based on improved SMOTE unbalanced sample processing and IHPO-DBN[J]. Power System Protection and Control, 2024, 52(11): 21-30.

[4] 康佳宇，张沈习，张庆平，等. 基于 ANOVA 和 BO-SVM 的变压器故障诊断方法[J]. 高电压技术，2023，49(5)：1882-1891.
KANG Jiayu, ZHANG Shenxi, ZHANG Qingping, et al. Fault diagnosis method of transformer based on ANOVA and BO-SVM[J]. High Voltage Engineering, 2023, 49(5): 1882-1891.

[5] 周远翔，陈健宁，赵亮，等. 电力设备数智化发展历程与大语言模型智能体应用展望[J]. 高电压技术，2025，51(8)：4089-4109.
ZHOU Yuanxiang, CHEN Jianning, ZHAO Liang, et al. Evolution of digitization and intelligence of power equipment and prospects for application of large language model based agent[J]. High Voltage Engineering, 2025, 51(8): 4048-4109.

[6] 谢庆，王春鑫，李帆，等. 知识及数据驱动的电力一次设备健康管理方法综述[J]. 高电压技术，2024，50(2)：605-620.
XIE Qing, WANG Chunxin, LI Fan, et al. Review of knowledge and data-driven primary equipment health management methods[J]. High Voltage Engineering, 2024, 50(2): 605-620.

[7] 唐文虎，牛哲文，赵柏宁，等. 数据驱动的人工智能技术在电力设备状态分析中的研究与应用[J]. 高电压技术，2020，46(9)：2985-2999.
TANG Wenhu, NIU Zhewen, ZHAO Boning, et al. Research and application of data-driven artificial intelligence technology for condition analysis of power equipment[J]. High Voltage Engineering, 2020, 46(9): 2985-2999.

[8] 李平，胡根铭. 基于数据增强型一维改进卷积神经网络的变压器故障诊断方法[J]. 电网技术，2023，47(7)：2957-2966.
LI Ping, HU Genming. Transformer fault diagnosis based on data enhanced one-dimensional improved convolutional neural network[J]. Power System Technology, 2023, 47(7): 2957-2966.

[9] HAN X H, HUANG S, MA S F, et al. Fault diagnosis methodfor transformer based on NCA and CapSA-RELM[J]. Electrical Engineering, 2024, 106(1): 203-213.

[10] 朱莉，汪小豪，李豪，等. 不平衡样本下基于变异麻雀搜索算法和改进 SMOTE 的变压器故障诊断方法[J]. 高电压技术，2023，49(12)：4993-5001.
ZHU Li, WANG Xiaohao, LI Hao, et al. Transformer fault diagnosis method based on variation sparrow search algorithm and improved SMOTE under unbalanced samples[J]. High Voltage Engineering, 2023, 49(12): 4993-5001.

[11] 谢国民，倪乐水，曹媛. 基于 VSRP 与 β-GWO-SVM 的变压器故障辨识方法[J]. 高电压技术，2021，47(10)：3635-3641.
XIE Guomin, NI Leshui, CAO Yuan. Transformer fault identification method based on VSRP and β-GWO-SVM[J]. High Voltage Engineering, 2021, 47(10): 3635-3641.

[12] 黄新波，蒋卫涛，朱永灿，等. 基于时间序列和支持向量机的变压器故障预测[J]. 高电压技术，2020，46(7)：2530-2538.
HUANG Xinbo, JIANG Weitao, ZHU Yongcan, et al. Transformer fault prediction based on time series and support vector machine[J]. High Voltage Engineering, 2020, 46(7): 2530-2538.

[13] HU S F, WU J, CIREN O, et al. Fault diagnosis of power transformers based on t-SNE and ECOC-TEWSO-SVM[J]. AIP Advances, 2024, 14(5): 055126.

[14] XUE J K, SHEN B. A novel swarm intelligence optimization approach: sparrow search algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.

[15] 王艳，李伟，赵洪山，等. 基于融合少数类过采样均衡多分类数据的改进极限学习机的变压器故障诊断方法[J]. 电网技术，2023，

47(9)：3799-3807.
WANG Yan, LI Wei, ZHAO Hongshan, et al. Transformer faultdiagnosis fused with synthetic minority over-sampling balanced multi-classification data based on improved extreme learning machine[J]. Power System Technology, 2023, 47(9): 3799-3807.

[16] DU X Z, CHEN K, DU H Y, et al. TS-SSA: an improved two-stage sparrow search algorithm for large-scale many-objective optimization problems[J]. PLoS One, 2025, 20(3): e0314584.

[17] 余松，胡东，唐超，等. 基于 TLR-ADASYN 平衡化数据集的 MSSA-SVM 变压器故障诊断[J]. 高电压技术，2021，47(11)：3845-3853.
YU Song, HU Dong, TANG Chao, et al. MSSA-SVM transformer fault diagnosis method based on TLR-ADASYN balanced data set[J]. High Voltage Engineering, 2025, 20(3): e0314584.

[18] 李艳霞，柴毅，胡友强，等. 不平衡数据分类方法综述[J]. 控制与决策，2019，34(4)：673-688.
LI Yanxia, CHAI Yi, HU Youqiang, et al. Review of imbalanced data classification methods[J]. Control and Decision, 2019, 34(4): 673-688.

[19] 崔宇，侯慧娟，苏磊，等. 考虑不平衡案例样本的电力变压器故障诊断方法[J]. 高电压技术，2020，46(1)：33-41.
CUI Yu, HOU Huijuan, SU Lei, et al. Fault diagnosis method for power transformer considering imbalanced class distribution[J]. High Voltage Engineering, 2020, 46(1): 33-41.

[20] GOODFELLOW I J. POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 28th Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2014: 2672-2680.

[21] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[C]//International Conference on Machine Learning. Sydney, Australia: JMLR, 2017: 214-223.

[22] 穆泽渊，蒋泽赟，王成龙，等. 变压器油中电弧的产气特性[J]. 高电压技术，2025，51(7)：3434-3442.
MU Zeyuan, JIANG Zeyun, WANG Chenglong, et al. Gas generation characteristics of arc in transformer oil[J]. High Voltage Engineering, 2025, 51(7): 3434-3442.

[23] 刘云鹏，许自强，和家慧，等. 基于条件式 Wasserstein 生成对抗网络的电力变压器故障样本增强技术[J]. 电网技术，2020，44(4)：1505-1513.
LIU Yunpeng, XU Ziqiang, HE Jiahui, et al. Data augmentation method for power transformer fault diagnosis based on conditional Wasserstein generative adversarial network[J]. Power System Technology, 2020, 44(4): 1505-1513.

[24] 陈嘉兴，刘扬，刘晓茜，等. Logistic 混沌映射与差分进化改进人工蜂群优化水下定位[J]. 工程科学与技术，2025，57(1)：57-67.
CHEN Jiaxing, LIU Yang, LIU Xiaoqian, et al. Improved artificial bee colony optimization underwater localization algorithm by Logistic chaos mapping and differential evolution[J]. Advanced Engineering Sciences, 2025, 57(1): 57-67.

[25] 李昕燃，靳伍银. 基于改进麻雀算法优化支持向量机的滚动轴承故障诊断研究[J]. 振动与冲击，2023，42(6)：106-114.
LI Xinran, JIN Wuyin. Fault diagnosis of rolling bearings based on ISSA-SVM[J]. Journal of Vibration and Shock, 2023, 42(6): 106-114.

[26] 赵靖英，吴晶晶，张雪辉，等. 基于萤火虫扰动麻雀搜索算法-极限学习机的光伏阵列故障诊断方法研究[J]. 电网技术，2023，47(4)：1612-1622.
ZHAO Jingying, WU Jingjing, ZHANG Xuehui, et al. Fault diagnosis of photovoltaic arrays based on sparrow search algorithmwith firefly perturbation-extreme learning machine[J]. Power System Technology, 2023, 47(4): 1612-1622.

[27] 田晓飞. 基于改进蝙蝠算法优化支持向量机的变压器故障诊断研究[D]. 成都：西华大学，2019.
TIAN Xiaofei. Research on transformer fault diagnosis based on improved bat algorithm optimizing support vector machine[D]. Chengdu, China: Xihua University, 2019.



曾贤强
1978—，男，博士，副教授，硕导
主要从事综合能源系统、变压器故障诊断等方面的工作
E-mail: zengxqlut@163.com

ZENG Xianqiang
Ph.D.
Associate professor



梁　嘉(通信作者)
2000—，男，硕士
主要研究方向为变压器故障诊断
E-mail: 1821656453@qq.com

LIANG Jia
Corresponding author

收稿日期　2025-07-15　在线出版日期　2026-01-29　　编辑　卫李静